TELEPHONE HISTORY ANALYSIS OF THE 2001 CALIFORNIA HEALTH INTERVIEW SURVEY

Amy Luo, Ernst and Young; Ismael Flores Cervantes, Westat; Charles DiSogra and Wei Yen, UCLA
Amy Luo, Ernst and Young, 1225 Connecticut Ave. NW, Washington, DC 20036

Key Words: **Nonresponse, telephone survey, call history, CHIS**

1. Introduction

Random digit dial (RDD) telephone surveys have become one of the most common methods for collecting data in socio-economic studies in recent years. Among the advantages of telephone surveys are reduced costs and shorter field period for data collection compared to in-person surveys. However, in recent years, decreasing response rates have become a major concern in these surveys. Special techniques and methods have been implemented to deal with this problem.

Low cooperation and response rates can weaken, or in worst cases, invalidate the main findings of a survey. According to the American Association for Public Opinion Research (AAPOR) 'Best Practice' guide, 'a low cooperation or response rate does more damage in rendering a survey's results questionable than a small sample.' (AAPOR, 1997) To ensure the validity of the findings in some government or academic surveys, there are explicit provisions that require a minimum response rate in the study (i.e., OMB has a 80% response rate requirement). Low response rates are also more likely to yield nonresponse errors. Groves (1989) modeled the nonresponse error as the product of nonresponse rate and the differences between nonrespondents and respondents. Although low nonresponse rates do not guarantee low nonresponse errors when there are distinct differences on the characteristics being measured between respondents and nonrespondents, high nonresponse rate will likely yield higher nonresponse errors.

Many survey organizations have taken efforts to deal with low response rates. Part of this effort focuses on procedures implemented after data collection. These procedures are mostly analytical and include the creation and use of nonresponse-adjusted weights or the implementation of imputation methods to compensate for both unit nonresponse and item nonresponse (Massy and Botman, 1988). However, organizations are focusing more effort on procedures that can be implemented before or during the data collection period. Those procedures include the use of incentives, pre-notification letters, questionnaire development, pilot surveys, optimal calling time scheduling, and refusal conversion protocols (Groves, 1983).

In this paper, we examine the optimal calling time during data collection, through the use of Computer-Assisted Telephone Interviewing (CATI). CATI systems are used in most large-scale surveys, and are the central tool for management of day-to-day data collection operations. Call scheduling, interview workload assignment and allocation of calls among the telephone centers are managed though CATI. Because CATI centralizes the control of the survey operations, functions such as daily monitoring and evaluation of interviewer and telephone center's performance are easily implemented.

This paper focuses on two contact protocols determined by the CATI scheduling system. The first protocol is the time of the first call attempt to the telephone number. The first call is made at a time when as many potential respondents as possible can be reached. However, calls during the optimal time cannot always be made because of limited resources (i.e., available equipment, and interviewers) or current workload. This situation is balanced throughout data collection until all telephone numbers in the sample are called for the first time. In this paper we will examine the time when the first call is made and if a respondent is reached in order to find optimal times for the first call.

The second contact protocol is at the refusal conversion stage, where potential respondents who have initially refused are contacted again. Specially trained interviewers call back initial refusals at a later time in order to seek their cooperation. This paper examines the time the survey organization should wait before calling back in order to obtain higher cooperation. In this paper, we refer to this time as "holding time", and we explore if there is an optimal holding time after the first refusal that would yield a higher refusal conversion rate.

There is ample literature that addresses the problem of the optimal time for the first telephone call attempt. Most studies found that weekends and weekday nights are usually good times to reach respondents (Wiseman and McDonald, 1979; Kerin and Peterson, 1983; Weeks et al., 1987). Weeks (1988) has the first detailed discussion of computerized call scheduling. In that study, he confirms that weekends and weekday nights are good time to call. In a more recent study, Brick et al. (1996) re-examines calling time patterns. His analysis considers socio-economic and geographic characteristics of the telephone exchange number thought to affect the initial call contact success. Although his results also confirm previous findings, his analysis suggests that there are differences in calling patterns for some demographic groups and geographic areas. The problem of the optimal first call time has not been revisited recently. Because of the increasing cellular telephone usage and changes in call privacy legislation in recent years, we re-examine if the

contact patterns have changed in anyway since they were studied.

In contrast, there is scant research on the optimal holding time for refusal conversion. Lessler and Kalsbeek (1992) suggest waiting a number of days before proceeding with refusal conversion attempts; however, there is no specific guidelines or recommendations on how long the waiting time should be in order to maximize the conversion of refusals. We present a predictive analysis on the yield of refusal conversion at different holding times in this paper.

The data for both analyses come from the call history file created by CATI for the 2001 California Health Interview Survey (CHIS 2001) sample. CATI systems maintained a file with the call history for each telephone number in the sample. The file includes day, time, telephone center, and the disposition code for each call attempt. The analysis in this paper focuses only on the contact protocols for the screening interview. We also used socio-economic and geographic information at the exchange level for telephones in the sample.

2. Survey data

CHIS 2001 was a collaborative project of the UCLA Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute, that focused on public health and access to health care. The survey was the largest state health survey ever undertaken in the United States. CHIS 2001 was a RDD telephone survey of California households designed to produce reliable estimates for the whole state, for large and medium-sized population counties in the state, and for groups of the smallest population counties. The survey also supported the study of the characteristics for the major racial and ethnic groups and a number of smaller ethnic groups within the state. Adults, parents or guardians of children, and adolescents within California households responded. Nearly 58,000 respondents were interviewed for CHIS 2001.

3. Call history information

3.1 First call attempts

The data collection period was approximate one year, from November 2000 to November 2001. The call history file for CHIS 2001 contains 1,764,935 screener interview call attempts for 223,509 telephone numbers for the RDD sample[1]. An average of 7.90 calls per number were made in order to complete 84,051 screener interviews. The original RDD sample consisted of 365,308 telephone numbers and 141,799 (38.8%) numbers were not dialed because they were determined to be non-residential, non-working or business numbers or were sub-sampled (California Health Interview Survey, 2001).

Westat (a survey research company) collected data for CHIS from 6 telephone centers. Four centers were located in the East coast (Maryland, New Jersey; Pennsylvania), one in Colorado and one in California. The centers in California and Colorado in the Pacific and Mountain Time zones were useful for scheduling of evening interviewing in California. Forty five percent of the telephone calls were made from centers in California and Colorado.

Operation hours were from 9 AM to 9 PM in California local time during weekdays. On weekends, calls were made from 10 AM to 6 PM on Saturdays and 2 to 9 PM on Sundays.

3.2 Refusal conversion

After a respondent refused to cooperate in the survey, he was usually called again as part of refusal conversion protocol. There were 59,569 telephone numbers where somebody refused to complete the screener interview. Only 77 percent of the refusals (45,824 numbers) were scheduled for conversion. This set of numbers constitutes the data used in the second part of the analysis.

As part of the standard conversion procedure, CATI would have scheduled additional call attempts two weeks after refusal. However, the call history showed that the holding time for conversion was more than two weeks. The actual time varies from two weeks to more than five weeks. Among the reasons for the delay were the availability of interviewers and the release of new work (telephone numbers dialed for the first time). Another reason was the time needed for the mailing of a refusal conversion letter. As part of the sampling procedures, sampled telephone numbers are matched with addresses. Then pre-notification letters are mailed out to improve response rates. If the respondent refused to answer the survey when contacted the first time, then a refusal conversion letter is mailed out to same address It is expected that the second letter would improve the odds of successful conversion when the refusal was called back at a later time. In these “mailable” households, CATI scheduled conversion calls after the mailing of the refusal conversion letter, which was for some cases beyond the two-week period.

Table 3-1 shows the distribution of refusals by the time in weeks when the refusal conversion calls were made. The fact that refusal conversion calls were not made after two weeks enabled us to determine if there was

[1] CHIS 2001 sample consisted of an RDD sample and several race-ethnic list samples. The list samples were excluded in this analysis.

an optimal holding time with a higher rate of refusal conversion based on the observed data. Table 3-1 also shows the number of refusals that were converted.

Table 3-1. Distribution of Refusals and Converted Refusal by Holding Time Period Before Refusal Conversion Calls.

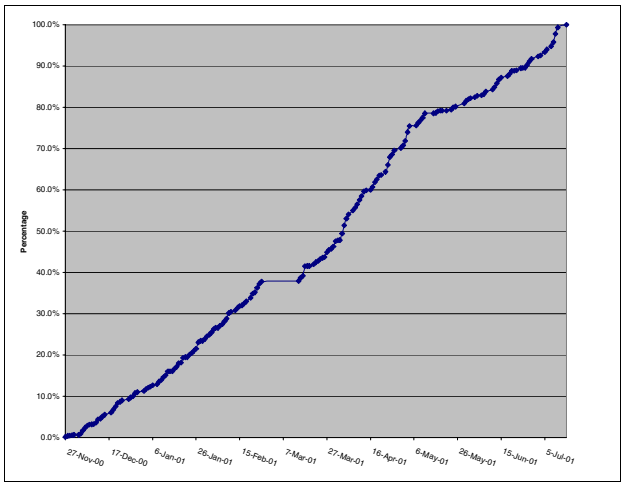
	Total Refusals		Converted Refusals	
Holding Time	Number	Percentage	Number	Percentage
2 weeks	6,934	15.1%	1,946	12.7%
2 -3 weeks	22,834	49.8%	8,061	48.5%
3-4 weeks	11,603	25.3%	4,800	28.9%
4-5 weeks	4,453	9.7%	1,804	10.9%
Total	45,824	100.0%	16,611	100.0%

4. Findings

4.1 Optimal calling time for the first attempt contacts

The distribution of the first attempt calls (i.e., new work) was fairly uniform throughout November 2000 and July 2001 (Figure 1). The data collection period was longer than planned due to the additional time required for refusal conversion and additional attempts for telephone numbers released at the end of the new work queue.

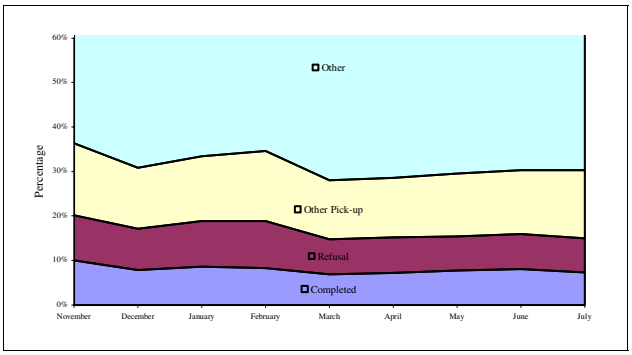
Figure 1. Cumulative Distribution of First Call Attempts in CHIS 2001.



We classified the first call outcome into the following four categories: (a) Completed, when somebody in the household answered the telephone and completed the interview. (b) Refusal, when somebody in the household answered the telephone, but refused to cooperate. (c) Other Pick-up, when somebody answered the telephone but neither refused nor completed the interview (i.e., scheduled a callback at a later time, or requested a mail out letter with additional information about the survey, etc.). (d) Other, that includes answer machine, non-working, busy signal, questionable ring and any other code for non-residential number. Answer machines were included in the 'other' category, because we are solely interested in optimal time to have a person pick up the phone.

We first look at the distribution of the results of the first contact by month. Because the data collection period lasted almost a year, we are able to see if the results of the first contact vary by months. Figure 2 shows the distribution of the results of the first contact for month.

Figure 2. Distribution of the First Attempt Calls by Months in CHIS 2001.

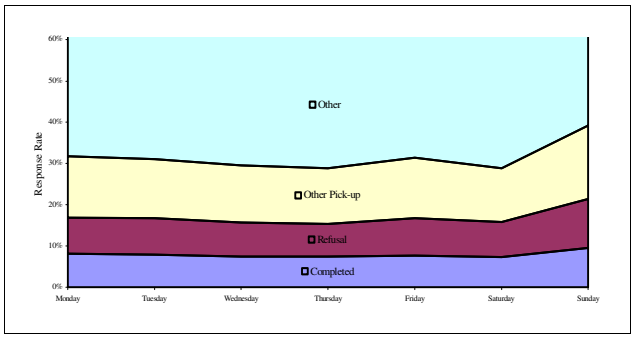


The test statistic indicates that there is a relationship between the first contact rate and the month when the first call attempt is made (χ^2=53141, df=8[2], p< 0.0001). The figure indicates that it is more likely to reach a respondent during wintertime, except for December. This is consistent with previous findings (Wiseman and McDonald, 1979; Kerin and Peterson, 1983; Weeks et al., 1987). November and February seem the two peak months in terms of contact rate. However, the refusal rate and completion rate are higher in November and February. In contrast, it is less likely to reach a respondent during spring and summer.

Figure 3 shows the first-call contact pattern by day of the week (χ^2= 32999, df =6, p < 0.0001). As previous research has shown, the probability of contacting a respondent is higher during weekends than weekdays. Although Sunday has the best contact rate, the likelihood of a successful contact is not as large as found in other studies.

Figure 3. Distribution of Results of First Call Attempts by Weekdays in CHIS 2001

[2] Most new telephone numbers were released from November to July, so the data includes a total of 9 months for the chi-square statistics



Because weekdays have different first contact patterns than weekends, we analyzed separately weekends and weekdays. Figure 4 shows the first-call results by time of the day (hour) during weekdays (χ^2=265696, df=11, $p < 0.0001$). Figure 5 shows the results for weekends and weekends (χ^2=28107, df=13, $p < 0.0001$). In weekdays, the likelihood of contacting a respondent is higher from 3:00 PM to 8:00 PM, similar to previous research (Brick et al., 1996; Odom and Kalsbeek, 2000; Weeks et al, 1987). However, our study suggests that late afternoon seems to have also high contact rates.

Figure 4. Distribution of Results of First Call Attempts by Weekday Hours in CHIS 2001

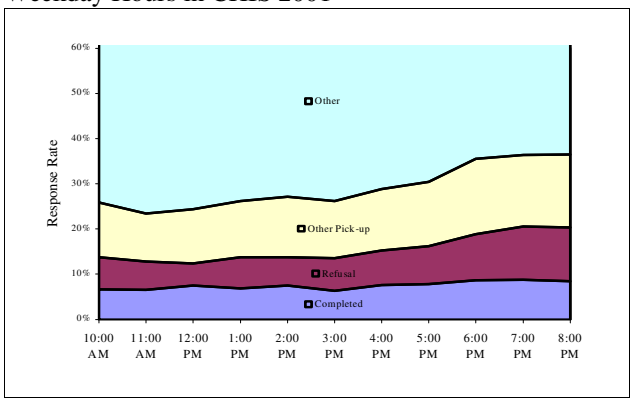
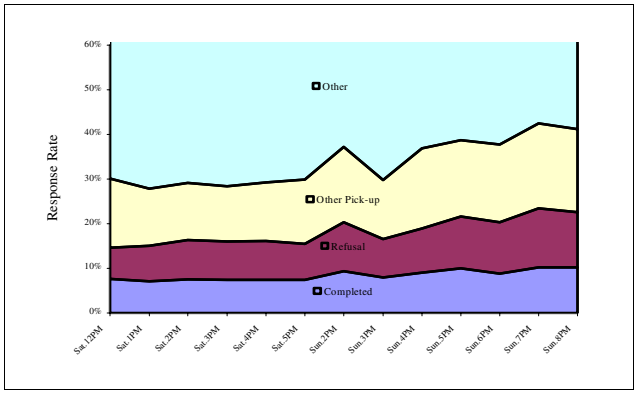


Figure 5. Distribution of Results of First Attempts Calls by Weekends Hours in CHIS 2001



4.2 Refusal Conversion

Preliminary analysis of the CHIS 2001 data shows a relationship between the refusal conversion rate and the holding time before any calls for conversion are made. Table 5-1 shows that refusals are more likely to be converted if conversion attempts are made after longer holding periods. As indicated in the table, the largest refusal conversion rate (41.4 percent) corresponds to refusal cases that are held for 3 to 4 weeks. This conversion rate is 1.5 times the observed conversion rate for 2 to 3 weeks (the default holding time used in CATI). These results imply the existence of an optimal holding time that would yield a higher number of refusal conversions. This optimal time could be used as part of the conversion protocols in future surveys. However, these observations should be studied before implementing any changes in the current CATI protocols.

Table 5-1. The Distribution of Observed Converted Refusal Cases by Holding Time

Holding time	Percentage of refusals converted	Ratio*
2 weeks	28.1 %	1.0
2-3 weeks	35.3 %	1.3
3-4 weeks	41.4 %	1.5
4-5 weeks	40.5 %	1.4
Total	36.2 %	

*with respect to the percentage of converted refusals called after two weeks.

In order to explore the holding time between a refusal conversion call and the first refusal that would yield a higher refusal conversion rate, a logistic regression model was built to predict the refusal conversion. The model is defined as

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i ,$$

where y_i is the binary dependent variable for whether a refusal is converted or not at the screener level, and x_i is a vector of predictor variables. There are two types of variables used in the regression model. One type is the CATI scheduling related variables, such as number of calls, or if a refusal conversion letter was mailed. The other type of predictor is the exchange level demographic variables, such as percentage of Hispanics, household renter/owner or persons 65 years old or older in the area covered by exchange of the telephone number.

The first model (or simple model) includes only the holding-time variable. Additional explanatory variables believed to be possible predictors for refusal conversion (demographic and CATI variables) are included in a

step-wise fashion. The inclusion of a new predictor in the model is evaluated by examining the goodness of fit, R^2 and additional diagnostics produced for the model. The most parsimonious model is selected at the end of the analysis.

Table 5-2. The Logistic Regression Models for Predicting the Successful Conversion of a First Time Refusal

	Simple Model		Full Model	
Predictors	**Odds Ratio**	**95% Wald C.L.**	**Odds Ratio**	**95% Wald C.L.**
Holding-time				
2 weeks	1		1	
2-3 weeks	1.398	(1.318, 1.483) *	1.003	(0.933, 1.079) *
3-4 weeks	1.808	(1.696, 1.928) *	1.189	(1.102, 1.284) *
4-5 weeks	1.745	(1.612, 1.890) *	1.106	(1.010, 1.211)
Demographics				
Percent Hispanic	\	\ \ \	0.992	(0.987, 0.997) *
Percent Rent	\	\ \ \	0.994	(0.993, 0.996) *
Percent 65 years old and up	\	\ \ \	1.003	(1.001, 1.004) *
MSA status				
1	\	\ \ \	1.000	
2	\	\ \ \	0.975	(0.905, 1.051) *
3	\	\ \ \	1.084	(1.025, 1.146)
4	\	\ \ \	1.189	(1.116, 1.266) *
5	\	\ \ \	1.154	(1.080, 1.233) *
TRC variables				
Mailed refusal conversion letter	\	\ \ \	1.695	(1.594, 1.802) *
Telephone center	\	\ \ \	1.000	
B	\	\ \ \	1.276	(1.160, 1.403) *
C	\	\ \ \	0.943	(0.871, 1.021) *
D	\	\ \ \	1.008	(0.922, 1.102) *
E	\	\ \ \	1.373	(1.273, 1.482) *
F	\	\ \ \	1.241	(1.138, 1.352) *
Initial number of calls			0.962	(0.957, 0.967) *

* significant at the 0.01 level or less

Table 5-2 shows the details of the simple and full model for this analysis. The table presents the logistic regression models, including the odds ratios and their associated confidence interval of the significant variables in the model. The first level or categorical value of the variables in the models is used as reference when computing the odds ratios.

In the simple model, the levels of the variable for holding time (2 to 3 weeks, 3 to 4 weeks, etc.) are significant at the 0.01 level. The model indicates that holding for 3 to 4 weeks after the refusal is 1.8 times more likely to yield a successful conversion when compared to the likelihood of conversion when holding for 2 weeks. The results of the simple model corroborate the observations made in Table 5-1.

However, when introducing additional explanatory variables in the model, the effect of the holding time variable is greatly diminished but not removed from the model. In these models, whether a refusal has an address that can be used for mailing out a letter has the largest effect on the refusal conversion outcome. The model shows that refusals with mailing addresses (i.e., "mailable" refusal) are much more likely to be converted than refusals without a mailing address ("non-mailable" refusal).

Further analysis includes separate models for mailable and non-mailable refusals. These models show that there is no relationship between the holding time and the refusal conversion rate for non-mailable refusals. In contrast, there is a holding time effect, which cannot be explained by any other available predictor for mailable refusals, but this effect is not as strong as initially seen. Although still significant, holding 3 to 4 weeks is only 1.3 times more likely to yield a successful conversion when compared to the likelihood of converting a refusal after holding for 2 weeks. However, these findings may be confounded by the time needed to allow for the refusal to read the conversion letter (i.e., a higher likelihood that the respondent has read the refusal conversion letter after 3 to 4 weeks rather than after 2 weeks). The analysis suggests that the optimal holding time is 3 to 4 weeks for mailable refusals.

5. Implications and Discussion

This paper presents the results of two analyses of specific protocols of the interview process that deal with nonresponse. The first analysis focuses on the optimal time of first call (i.e., month, day of the week, and time of the day) that is most likely to reach a respondent in the household. The second analysis focuses on the time the interviewer should wait before making a refusal conversion call that is most likely to yield a successful conversion.

These analyses have their limitations and the results should not be considered as part of the development of optimal calling protocols that could be used for all CATI surveys. These results are also restricted to California residents and do not reflect patterns found in other areas of

the country. CHIS 2001 had a long data collection period not common in other RDD surveys. Nevertheless, these results provide valuable insights to the factors to be considered in future cycles of CHIS so higher response rates can be achieved in a more cost-effective way. The main findings of this study are listed below:

1) First call contact rates throughout the year. The findings of this paper are consistent with previous studies as reported in the published literatures. The first call contact rate varies throughout the year. The contact rate in winter is generally higher than the rate in spring and summer. Although December has the lowest contact rate among the winter months, the December rate is still higher than the rate in spring. A possible explanation may be that that people are more likely to be at home in winter, therefore, easier to reach than summer, spring or during the holidays.

2) First call contact rates throughout the days of the week and time of the day. Again, the findings confirm the observations from previous studies. There are significant differences in contact rates by day of the week and time of the day. In general, weekends are better than weekdays, and weekday nights are better than daytime weekdays. In addition, weekday late afternoons are good times to reach potential respondents. During weekends, Sundays are better than Saturdays. Sunday night is the best time during the week. Depending on time and workload, there are potential gains if more calls are made on Sundays, in particular, Sunday nights.

3) Holding time periods. Although simple tabulations of the data suggest that longer holding periods yield higher rates of refusal conversion, the net effect is not as strong as initially thought once the effect of other explanatory variables are accounted for. The holding time effect does not exist for refusals who do not receive a conversion letter, and this effect is small for refusals who receive it. Nevertheless, the analysis shows that holding for 3 to 4 weeks would yield the highest conversion rate compared with other holding times. Nevertheless, these results are not conclusive. It is recommended to conduct an experiment by randomly assigning the holding time to telephone numbers before conversion in future rounds of CHIS in order to confirm the results of this study.

4) Refusal conversion letter. The analysis shows that the effect of the conversion letter is more important than the effect of the holding time. This effect is similar to the effect on the response rates of pre-notification letters sent before the respondent is contacted for the first time. As found in previous studies, there are potential gains if sampled telephone numbers are matched to mailing addresses. In this way, response rates can be improved by sending pre-notification and refusal conversion letters to theses addresses.

6. References

Brick, J.M, Allen, B., Cunningham, P., and Maklan D. (1996), "Outcomes of a Calling Protocol in a Telephone Survey", 1996 *Proceedings of the American Statistical Association,* Survey Research Methods Section, Alexandria, VA: American Statistical Association: 142-149.

California Health Interview Survey (2001) Weighting and Variance Estimation *Methodology Series: Report 5 –*. Los Angeles, CA: UCLA Center for Health Policy Research.

Curtin, R., Presser, S. and Singer, E. (2000), "The Effects of Response Rate Changes on the Index of Consumer Sentiment", *Public Opinion Quarterly*, 64, 413-428.

Groves, R. M. (1989), *Survey Errors and Survey Costs*, New York: Wiley.

Groves, R. M., Biemer, P.P., Lyberg, L. L., Massey, J.T., Nicholls, W. L., and Waksberg, J. (1988), (eds.) *Telephone Survey Methodology*. New York: John Wiley & Sons.

Keeter, S., Kohut, A., Miller, C., Groves, R.M. and Presser, S. (2000), "Consequences of Reducing Nonresponse in a Large National Telephone Survey", *Public Opinion Quarterly*, 64, 125-148

Kerin, R. and Peterson, R. (1983), "Scheduling Telephone Interviews: Lessons From 250,000 Dialing", *Journal of Advertising Research*, 23, 41-47.

Massey, J.T., and Botman, S.L. (1988). "Weighting Adjustments for Random Digit Dialed Surveys," in R.M. Groves, P.P. Biemer, L.L. Lyberg, J.T. Massey, W.L. Nicholls, and J. Waksberg (eds.). *Telephone Survey Methodology*. New York: John Wiley & Sons, 143-160.

Odom, D. M. and Kalsbeek, W.D. (2000), "Further Analysis of Telephone Call History Data from the Behavioral Risk Factor Surveillance System", *2000 Proceedings of the American Statistical Association,* Survey Research Methods Section, Alexandria, VA: American Statistical Association: 398-403.

O'Neil, M. J. (1979), "Estimating the Nonresponse Bias Due to Refusals in Telephone Surveys", *Public Opinion Quarterly*, 43, 2, 218-232.

Potthoff, R.F., K. G. Manton, and M. A. Woodbury (1993), "Correcting for Nonavailability Bias in Surveys by Weighting Based on Number of Callbacks (in Applications and Case Studies)", *Journal of the American Statistical Association*, 88, 424, 1197-1207.

The American Association for Public Opinion Research (1997), *Best Practices for Survey and Public Opinion Research.*

Weeks, M. F. (1988), "Call Scheduling with CATI: Current Capabilities and Methods", in R. M. Groves, P. P. Biemer, L.E. Lyberg, J. Massey, W. L. Nicholls II, and J. Waksberg (eds.), *Telephone Survey Methodology*, New York: Wiley, 403-420.

Weeks, M. F., B. L. Jones, R. E. Folsom, Jr., C. H. Benrud (1980), "Optimal Times to Contact Sample Households", *Public Opinion Quarterly*, 44, 1, 101-114.

Weeks, M. F., R. A. Kulka, and S. A. Pierson (1987), "Optimal Call Scheduling for a Telephone Survey", *Public Opinion Quarterly*, 51, 4, 540-549.

Wiseman and McDonald, (1979), "Noncontact and Refusal Rates in Consumer Telephone Surveys", *Journal of Marketing Research*, 16, 478-484.